ゼミ資料 (Distributed Estimation)

藤田研究室　　　山田　照樹

平成 18 年 6 月 5 日

1　はじめに

本レポートは, Calfornia Institute of Technology (caltech) の R. Murray らによる Networked Control System という講義の 7-1 の資料 (Distributed Estimation) についてまとめる．前半は, 2 乗誤差の期待値を評価関数とし, それを最小化する最小分散推定値を一般的な形で表わし, いくつかのケースを例に考えている．また, Static sensor fusion として, 個々のセンサで得られるローカルな推定値に重みをかけたものを統合することによりグローバルな推定値が得られるとしている．後半は, Dynamic sensor fusion として, 個々のセンサが Kalman Filter のアルゴリズムによりローカルな推定値と推定誤差の共分散を求めることにより, データの統合をシンプルにおこなえるようになることをみる．すなわち, 分散的に推定をおこなうことにより, 計算の負荷がひとつのノードに集中しなくなる.

2　Preliminaries

2.1　Optimal mean square estimate of a random variables

Y は 他のランダムな変数 X に従属なランダムな変数とする．このとき, 2 乗誤差 (mean square error) の期待値 $E[X-\hat{X}]^2$ を最小化するような推定値 $\hat{X}$ を考える.

Proposition 2
上記の前提の下では, 最小分散推定値 (minimum mean square estimate) は, $E[X|Y=y]$ である.
ただし, $E[X|Y=y]$ は, $Y=y$ の下での x の条件つき期待値 (conditional expectation) である.

証明 1 推定値を表わす関数を $g(Y)$ とする．また, $f_{X,Y}(x,y)$ を X,Y の結合確率密度関数 *(joint probability density function)* とする．このとき, 2 乗誤差の期待値は

$$\begin{aligned} E[X-\hat{X}]^2 &= \int_x \int_y (x-g(y))^2 f_{X,Y}(x,y)dxdy \\ &= \int_y dy f_Y(y) \int_x (x-g(y))^2 f_{X|Y}(x|y)dx \end{aligned}$$

ここで, $f_{X,Y}(x,y)=f_Y(y)f_{X|Y}(x|Y)$ を用いた．ただし, $f_{X|Y}$ は Y に関する X の条件つき確率密度関数とする．この値を最小化する推定値を求めたいので, 上の式を $g(y)$ で編微分すると

$$\begin{aligned} \frac{\partial E(X-\hat{X})^2}{\partial g(y)} &= -\int_y dy f_Y(y) \int_x 2(x-g(y)) f_{X|Y}(x|y)dx \\ &= 2\left(\int_y dy f_Y(y)(g(y) - \int_x x f_{x|Y}(x|y)dx)\right) \\ &= 2\int_y dy f_Y(y)(g(y) - E[X|Y=y]) \end{aligned} \tag{2.1}$$

となる．したがって, Y に関する X の条件つき期待値 $g(y)=E[x|Y=y]$ において 2 乗誤差の期待値が最小化される.

この結果は, X, Y がランダムなベクトル変数に対しても成り立つ．また, ランダムな変数がガウス分布に従う (ガウシアン変数) 場合には, 2 乗誤差の期待値を最小化する最小分散推定値は, Y について線形になることが知られている．ここからは, すべてのランダムな変数の平均値は 0 であるとする．また, ランダムな変数 X の共分散 (covariance) を R_x, X と Y の相互共分散 (cross-covariance) を R_{XY} と表わす.

Proposition 3

Y=y のとき, 2 乗誤差を最小化する X の最小分散推定値は

$$\hat{x} = R_{XY} R_Y^{-1} y \tag{2.2}$$

である．このとき, 推定誤差共分散は

$$P = R_X - R_{XY} R_Y^{-1} R_{YX} \tag{2.3}$$

である.

証明 2 最小分散推定値は y に関して線形であるので, その推定値を $\hat{x} = Ky$ とおくと, 推定誤差共分散行列 C は

$$\begin{aligned} C &= E[(x - Ky)(x - Ky)^T] \\ &= R_X - KR_{XY} - R_{XY}K^T + KR_YK^T \end{aligned} \tag{2.4}$$

C を K に関して編微分をおこない, C を最小化する Kを求めると

$$-2R_{XY} + 2KR_Y = 0 \Leftrightarrow K = R_{XY} R_Y^{-1} \tag{2.5}$$

となる．このとき, C は 最適な推定誤差共分散 P は, (2.4) 式にいま得られた K を代入すると

$$\begin{aligned} P &= R_X - R_{XY} R_Y^{-1} R_{YX} - R_{XY} R_Y^{-T} R_{XY}^T + R_{XY} R_Y^{-1} R_Y R_Y^{-T} R_{XY}^T \\ &= R_X - R_{XY} R_Y^{-1} R_{YX} \end{aligned} \tag{2.6}$$

と求まる.

Proposition 4

$y = Hx + v$ とする．H はある行列, v は平均が 0 で共分散行列が R_V のガウシアンノイズとする．また, v は X と独立とする．Y=y のもとで, 2 乗誤差を最小化する X のの最小分散推定値 $\hat{x}$ は

$$\hat{x} = R_X H^T (HR_X H^T + R_V)^{-1} y \tag{2.7}$$

で表わされ, このときの最適な推定誤差共分散行列 P は

$$P = R_X - R_X H^T (HR_X H^T + R_V)^{-1} HR_X \tag{2.8}$$

証明 3 $y = Hx + v$ のとき, Xと Yの相互共分散 R_{XY}, Y の共分散 R_Y を計算し, *Proposition 2*で求まった 最小分散推定値の式 (2.2) により求める.

$$\begin{aligned} R_{XY} &= E[XY^T] \\ &= E[x(Hx + v)^T] \\ &= R_X H^T \ \ (\because v \text{が} x \text{と独立な変数より}, Exv^T = 0) \end{aligned} \tag{2.9}$$

また, Y の共分散 R_Y は

$$\begin{aligned} R_Y &= E[YY^T] \\ &= E[(Hx + v)(Hx + v)^T] \\ &= E[Hxx^T H^T] + E[vv^T] \\ &= HR_X H^T + R_v \end{aligned} \tag{2.10}$$

と求まる．(2.9), (2.10) を (2.2), (2.3) へ代入することにより (2.7), (2.8) を得る.

3 Combining Estimators: Static sensor Fusion

この節では, 前節で求められた最小分散推定値が出力に関して線形な形で与えられていたものを別の形式に書き表わす.

Proposition 5

$y = Hx + v$ とする. このとき, 前節の Proposition 4 の最小分散推定値を別の形式で表わせば

$$P^{-1}\hat{x} = H^T R_V^{-1} y \tag{3.1}$$

となる. ここで, P は最適な推定誤差共分散行列であり

$$P^{-1} = (R_X^{-1} + H^T R_V^{-1} H) \tag{3.2}$$

と書き表せる.

このような形式に推定値を書き換えることによって, 推定値は, 各センサごとの推定値を組み合わせたものとして扱うことができ, すべてのセンサの測定値を中央の推定器に送って大規模な推定をおこなわせる必要はなくなる. このように, 各センサにおいてローカルな推定をおこない, それを中央のデータ処理部 (central data processing unit) へ送られ, そこで全体の推定値が求められる. これは static sensor fusion と呼ばれる.

3.1 Static Sensor Fusion for Star topology

Proposition 6 X をランダムな変数とし, それぞれのセンサの測定値 (measurement)y_i を基に推定される.

$$y_i = H_i x + v_i, \;\; i = 1, ..., n \tag{3.3}$$

ここで, v_i はセンサの測定値に混入するノイズであり, 他のセンサにおけるノイズ と無相関 (uncorrelated) であり, X とも無相関である. n 個の測定値に基づく x の推定値を $\hat{x}$, ひとつのセンサの測定値 y_i に基づく x の推定値を $\hat{x}_i$ と表わす. このとき, $\hat{x}$ は

$$P^{-1}\hat{x} = \sum_{i=1}^{n} P_i^{-1}\hat{x}_i \tag{3.4}$$

のように 各センサの測定値 y_i により求まる推定値 $\hat{x}_i$ の和の形で表わせる.

また, P は, $\hat{x}$ に対する推定誤差共分散行列であり, P_i は, $\hat{x}_i$ に対する推定誤差共分散行列であるとすると,

$$P^{-1} = \sum_{i=1}^{n} P_i^{-1} - (n-1)R_X^{-1} \tag{3.5}$$

と書ける.

証明 4 *y は, 各センサの測定値 y_i を縦に集めたベクトルとする. H, v に関しても同様に定義する. グローバルな推定値 (global estimate) $\hat{x}$ は*

$$P^{-1}\hat{x} = H^T R_V^{-1} y \tag{3.6}$$

のようにもとまることは, proposition 5 でみた. v_i はセンサごとお互いに無相関であるから, R_V は, 対角ブロックに v_i の共分散行列である R_{V_i} を並べた行列となる. よって, (3.6) 式の右辺は,

$$H^T R_V^{-1} y = \sum_{i=1}^{n} H_i^T R_{V_i} y_i \tag{3.7}$$

のように各センサの測定値に基づく形で書ける. ところで, $H_i^T R_{V_i} y_i$ はローカルな推定値 $\hat{x}_i$ を用いて

$$P_i^{-1}\hat{x}_i = H_i^T R_{V_i}^{-1} y_i \tag{3.8}$$

であるから,

$$P^{-1}\hat{x} = \sum_{i=1}^{n} P_i^{-1}\hat{x}_i \tag{3.9}$$

とできる．また

$$\begin{aligned} P^{-1} &= (R_X^{-1} + H^T R_V^{-1} H) \\ &= (R_X^{-1} + \sum_{i=1}^{n} H_i^T R_{V_i} H_i) \\ &= (R_X^{-1} + \sum_{i=1}^{n} (P_i - R_X^{-1})) \\ &= \sum_{i=1}^{n} P_i - (n-1) R_X^{-1} \end{aligned} \tag{3.10}$$

となる．

これにより, データを統合するセンター (fusion center) における計算の複雑さが低減される．また, ローカルな推定値は, 推定誤差共分散の逆行列の重み付けをされるので, 推定誤差の小さいローカルな推定値により大きな重みを付けることになり, その推定値により信頼性を置くことになる．

3.2 Static Sensor Fusion for Arbitrary Graphs

前節までで, 他のすべてのノードと通信ができ, データを統合する中心となるノード (central node) を仮定していた．まず, ローカルな推定値が中心のノードへ送られる．中心のノードは, ローカルな推定値にそれぞれ重み付けしてその合計を周囲のノードへ送り返す．実際には, 各センサは, 重みつきの推定値の平均値が必要となるので, この節では, 任意のグラフを考えて, average consensus 問題としてこれを捉えなおす．ただし, グラフは時不変で連結であるとする．このとき, 各ノードの状態変数のダイナミクスとして

$$x_i(k+1) = x_i(k) + h \sum_{j\ i} (x_j(k) - x_i(k)) \tag{3.11}$$

を考える．ここで, $j\ i$ は, ノード i がノード j につながれていることを示す．また, h はある小さな正の定数とする．これをすべてのノードについてまとめると,

$$x(k+1) = (I - hL)x(k) \tag{3.12}$$

となる．グラフラプラシアン L は, グラフが連結であるとき, 講義ノートに書かれている 3 つの特性をもち, この特性から average consensus が達成される．これを, weighted average consensus 問題へ拡張する．各ノードのダイナミクスは

$$x_i(k+1) = x_i(k) + hW_i^{-1} \sum_{j\ i} (x_j(k) - x_i(k)) \tag{3.13}$$

のように書ける．ただし, W_i は, x_i に対する正定な重みである．文献 [1] では, (3.13) を繰り返すことにより, すべての状態が

$$x_i(k) = x_j(k) = \frac{\sum_{i\in V} W_i x_i(0)}{\sum_{i\in V} W_i} \quad (i \neq j) \tag{3.14}$$

へ収束することが示されている．ここで, V はすべてのノードの集合を表わす．したがって, $x_i(0)$ をローカルな推定値 $\hat{x}_i$, W_i を $\hat{x}_i$ の推定誤差共分散の逆数 P_i^{-1} とみれば, weighted consensus が達成されることで, すべてのノードに重み付き推定値の和が伝わる．

3.3 Sequential Measurement from one sensor

この節では, ひとつのセンサから複数の測定値 (過去から現在までの測定値) が得られる場合において, 前節までのアルゴリズムを拡張することを考える. つぎのようなダイナミクスをもつランダムな状態変数 $x(k)$ を考える.

$$x(k+1) = Ax(k) + w(k) \tag{3.15}$$

$w(k)$ を平均が 0, 共分散行列が Q のガウス白色雑音とする. センサは, ステップ時間おきに測定値を更新し,

$$y(k) = Cx(k) + v(k) \tag{3.16}$$

とする. $v(k)$ を平均が 0, 共分散行列が R のガウス白色雑音とする. ここで, すべての測定値 $y(0), y(1), ..., y(k))$ が与えられている場合にグローバルな推定値を求める問題を考える. 与えられている測定値をつぎのような二つの集合に分割する.

1. 現在の測定値 $y(k)$
2. 過去の測定値 $y(0), ...y(k-1)$ の集合 Y

まずはじめに, $y(k)$ に基づく推定値 $\hat{x}$ をもとめると, (3.1) 式より

$$M^{-1}\hat{x} = C^T R^{-1} y(k) \tag{3.17}$$

となる. ここで, 推定誤差共分散行列 M は (3.2) 式より

$$M^{-1} = R_{x(k)}^{-1} + C^T R^{-1} C \tag{3.18}$$

である. $\hat{x}(k-1|k-1)$ を 先程定義した 2 の過去の測定値集合 Y に基づく $x(k-1)$ の推定値とし, $P(k-1|k-1)$ を その推定値に対応する推定誤差共分散行列を $P(k-1|k-1)$ と表わす. また, Y に基づく $x(k)$ の推定値 $\hat{x}|k-1$ は

$$\hat{x}(k|k-1) = A\hat{x}(k-1|k-1) \tag{3.19}$$

という関係式が成り立つ. これは $x(k) = Ax(k-1) + v(k-1)$ の両辺を Y に対する条件つき期待値をとることによって示される. また, $\hat{x}(k|k-1)$ に対応する推定誤差共分散行列 $P(k|k-1)$ は

$$\begin{aligned}
P(k|k-1) &= E[(x(k) - \hat{x}_{(k|k-1)})((x(k) - \hat{x}_{(k|k-1)})^T] \\
&= E[(Ax(k-1) + w(k-1) - A\hat{x}_{(k-1|k-1)})(Ax(k-1) + w(k-1) - A\hat{x}_{(k-1|k-1)})^T] \\
&= E[(A(x(k-1) - \hat{x}_{(k-1|k-1)}) + w(k-1))(A(x(k-1) - \hat{x}_{(k-1|k-1)}) + w(k-1))^T] \\
&= AP(k-1|k-1)A^T + Q
\end{aligned} \tag{3.20}$$

と求まる. ここで, ノイズ $w(k-1)$ が x に無相関であることを用いた. また, $x(k)$ のすべての測定値による推定値は, ローカルな推定値の結合によって求められる. Y に基づくローカルな推定値は $P(k|k-1)^{-1}\hat{x}(k|k-1)$ であり, $y(k)$ に基づく推定値は $M^{-1}\hat{x}(k)$ であるから, すべての測定値に基づく推定値 $P(k|k)^{-1}\hat{x}(k|k)$ は, これらの和として

$$P(k|k)^{-1}\hat{x}(k|k) = P(k|k-1)^{-1}\hat{x}(k|k-1) + M^{-1}\hat{x}(k) = P(k|k-1)^{-1}\hat{x}(k|k-1) + C^T R^{-1} y(k) \tag{3.21}$$

と書ける. 対応する推定誤差共分散行列は

$$P(k|k)^{-1} = P(k|k-1)^{-1} + C^T R^{-1} C \tag{3.22}$$

と表わせる. これらの関係式は, 時間と測定値がステップ時間おきに更新する Kalman Filter のアルゴリズムになっている.

4 Combining Sequential Measurement from Multiple Sensors: Dynamic sensor Fusion

この節では, 複数のセンサから得られる測定値を中央のノードに送り, そこでデータを統合し推定値を求める際に問題となることとそれに対する対応策を提案している. 中央のノードは, Kalman filter のアルゴリズムにより状態推定をおこなう. しかし, この手法にはつぎのような問題点がある.

1. 中央のノードは, センサーの数 (ノードの数) が多くなるほどサイズの大きい行列を扱って計算をおこなう必要がある.
2. ステップ時間ごとにセンサから測定値を送れるとは限らない.
ここでは, センサはステップ時間おきにデータを送れるものとして, 対策として中央のノードに計算の負荷を軽減させることを考えている.

5 Transmitting Local Estimates

この節では, 各センサが現時刻までの測定値を基にローカルな状態推定をおこない, それを中央のノードに送り結合することで, グローバルな推定値を得られるかどうかについて検討されている. しかし, それは容易ではないことが例を用いて示されている.

6 Distributed Kalman Filtering

information form of the Kalman filter

$$P^{-1}(k|k)\hat{x}(k|k) = P^{-1}(k|k-1)\hat{x}(k|k-1) + C^T R^{-1} y(k) \tag{6.1}$$

$$P^{-1}(k|k) = P^{-1}(k|k-1) + C^T R^{-1} C \tag{6.2}$$

ステップ時間ごとに $x(k|k)$ を求めていくためには

$$\hat{x}(k|k-1) = A\hat{x}(k-1|k-1) \tag{6.3}$$

$$P(k|k-1) = AP(k-1|k-1)A^T + Q \tag{6.4}$$

も必要となる. **Proposition 9** グローバルな推定誤差共分散行列 $P^{-1}(k|k)$ とグローバルな推定値 $\hat{x}(k|k)$ は, N 個のセンサから得られるローカルな推定誤差共分散行列 $P_i(k|k), P(k|k-1)$ とローカルな推定値 $\hat{x}_i(k|k), \hat{x}_i(k|k-1)$ を用いて

$$P^{-1}(k|k) = P^{-1}(k|k-1) + \sum_{i=1}^{N}(P_i^{-1}(k|k) - P_i^{-1}(k|k-1))$$

$$P^{-1}(k|k)\hat{x}(k|k) = P^{-1}(k|k-1)\hat{x}(k|k-1) + \sum_{i=1}^{N}(P_i^{-1}(k|k)\hat{x}_i(k|k) - P_i^{-1}(k|k-1)\hat{x}_i(k|k-1)) \tag{6.5}$$

のように表わせる.
dynamic sensor fusion に関して 2 つの方法が提案がされている.
1. 個々のセンサがローカルな推定値 $\hat{x}_i(k|k)$ を global fusion center へ送り, Proposition 9 により, グローバルな推定値を求める. ここで用いられている $\hat{x}(k|k-1), \hat{x}_i(k|k-1)$ は fusion ノードにより

$$\hat{x}(k|k-1) = A\hat{x}(k-1|k-1) \tag{6.6}$$

のように計算される. 同様に, すべての推定誤差共分散行列は, センサノードから他にデータが送られなくても, 計算できる. この手法はシンプルであるが, fusion ノードに計算の負担が大きいという問題がある.

2. センサノードから送られるデータ量が多くなるが, fusion ノードの計算の負担を軽減させる方法として, 式 (6.5) の右辺の $P^{-1}(k|k-1)\hat{x}(k|k-1)$ という項をつぎのような形で表わし, この情報もローカルなデータとして, 各センサから fusion node へ送る.

$$P^{-1}(k|k-1)\hat{x}(k|k-1) = \sum_{i=1}^{N} z_i(k) \tag{6.7}$$

証明 5 ここでは, $P^{-1}(k|k-1)\hat{x}(k|k-1)$ がローカルなデータの和として (6.7) 式のように書き表せることを示す. 代数計算により

$$\begin{aligned}
P^{-1}(k|k-1)\hat{x}(k|k-1) =& P^{-1}(k|k-1)A\hat{x}(k-1|k-1) \\
=& P^{-1}(k|k-1)AP(k-1|k-1)P^{-1}(k-1|k-1)\hat{x}(k-1|k-1) \\
=& P^{-1}(k|k-1)AP(k-1|k-1)(P^{-1}(k-1|k-2)\hat{x}(k-1|k-2) \\
&+ \sum_{i=1}^{N}(P_i^{-1}(k-1|k-1)\hat{x}_i(k-1|k-1) - P_i^{-1}(k-1|k-2)\hat{x}_i(k-1|k-2)))
\end{aligned} \tag{6.8}$$

したがって

$$z_i(k) = P^{-1}(k|k-1)AP(k-1|k-1)\left(z_i(k-1) + (P_i^{-1}(k-1|k-1)\hat{x}_i(k-1|k-1) - P_i^{-1}(k-1|k-2)\hat{x}_i(k-1|k-2))\right)$$

により, データは更新される.
以上から

$$\begin{aligned}
P^{-1}(k|k)\hat{x}(k|k) &= P^{-1}(k|k-1)\hat{x}(k|k-1) + \sum_{i=1}^{N}(P_i^{-1}(k|k)\hat{x}_i(k|k) - P_i^{-1}(k|k-1)\hat{x}_i(k|k-1)) \\
&= \sum_{i=1}^{N} z_i(k) + \sum_{i=1}^{N}(P_i^{-1}(k|k)\hat{x}_i(k|k) - P_i^{-1}(k|k-1)\hat{x}_i(k|k-1)) \\
&= \sum_{i=1}^{N}(z_i(k) + P_i^{-1}(k|k)\hat{x}_i(k|k) - P_i^{-1}(k|k-1)\hat{x}_i(k|k-1))
\end{aligned} \tag{6.9}$$

となる.

したがって, fusion ノードは, 各センサから送られてくるデータを足し合わせるだけでグローバルな推定値を求めることができる. よって, この方法は, 各センサから送られるデータ量は多くなるが, fusion ノードでの計算負荷は軽減されると考えられる.

7 おわりに

本レポートでは, caltech の Networked Control Systems の講義ノートの 7-1 をまとめた. 前半は, 2 乗誤差の期待値を最小化するという観点から, 最小分散推定値を求めた. つぎに, 各センサで得られるローカルな状態変数の推定値を fusion ノードに伝えて, データを統合することにより, グローバルな推定値を求める方法について紹介した. 最後に, 各センサに Kalman Filter を備えた (distributed Kalman Filter) のアルゴリズムを用いることにより, fusion ノードにおける計算量が軽減されることをみた.

参考文献

[1] D. Spanos and M. Murray, “Distributed Sensor Fusion using Dynamic Consensus,” *The 16th IFAC World Congress, Prague, Crech*, July. 2005.